==========
Sequence Listing was accepted.
If you need help call the Patent Electronic Business Center at (866) 217-9197 (toll free).
Reviewer: Anne Corrigan
Timestamp: [year=2008; month=12; day=10; hr=8; min=36; sec=12; ms=270;]
==========

Validated By CRFValidator v 1.0.3

Application No: 10591029 Version No: 1.0

Input Set:

Output Set:

Started: 2008-12-09 15:33:17.645
Finished: 2008-12-09 15:33:20.383
Elapsed: 0 hr(s) 0 min(s) 2 sec(s) 738 ms
Total Warnings: 49
Total Errors: 0
No. of SeqIDs Defined: 49
Actual SeqID Count: 49

Error code		Error Description
W	213	Artificial or Unknown found in <213> in SEQ ID (1)
W	213	Artificial or Unknown found in <213> in SEQ ID (2)
W	213	Artificial or Unknown found in <213> in SEQ ID (3)
W	213	Artificial or Unknown found in <213> in SEQ ID (4)
W	213	Artificial or Unknown found in <213> in SEQ ID (5)
W	213	Artificial or Unknown found in <213> in SEQ ID (6)
W	213	Artificial or Unknown found in <213> in SEQ ID (7)
W	213	Artificial or Unknown found in <213> in SEQ ID (8)
W	213	Artificial or Unknown found in <213> in SEQ ID (9)
W	213	Artificial or Unknown found in <213> in SEQ ID (10)
W	213	Artificial or Unknown found in <213> in SEQ ID (11)
W	213	Artificial or Unknown found in <213> in SEQ ID (12)
W	213	Artificial or Unknown found in <213> in SEQ ID (13)
W	213	Artificial or Unknown found in <213> in SEQ ID (14)
W	213	Artificial or Unknown found in <213> in SEQ ID (15)
W	213	Artificial or Unknown found in <213> in SEQ ID (16)
W	213	Artificial or Unknown found in <213> in SEQ ID (17)
W	213	Artificial or Unknown found in <213> in SEQ ID (18)
W	213	Artificial or Unknown found in <213> in SEQ ID (19)
W	213	Artificial or Unknown found in <213> in SEQ ID (20)

Input Set:

Output Set:

Started: 2008-12-09 15:33:17.645
Finished: 2008-12-09 15:33:20.383
Elapsed: 0 hr(s) 0 min(s) 2 sec(s) 738 ms
Total Warnings: 49
Total Errors: 0
No. of SeqIDs Defined: 49
Actual SeqID Count: 49

Error code	Error Description

This error has occured more than 20 times, will not be displayed

SEQUENCE LISTING

<110> WOOD, DAVID W.
HSII, JUDY
OAK, SEACHOL
CONTRERAS, LYDIA
CHESTNUT, JOHN

<120> SELF-CLEAVING AFFINITY TAGS AND METHODS OF USE

<130> 331772-00103

<140> 10591029
<141> 2008-12-09

<150> PCT/US05/05763
<151> 2005-02-24

<150> 60/548,092
<151> 2004-02-27

<160> 49

<170> PatentIn version 3.3

<210> 1
<211> 504
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
mini-intein derived from the full-length Mycobacterium
tuberculosis recA intein sequence

<400> 1
gccctcgcag agggcactcg gatcttcgat ccggtcaccg gtacaacgca tcgcatcgag 60

gatgttgtcg gtgggcgcaa gcctattcat gtcgtggctg ctgccaagga cggaacgctg 120

catgcgcggc ccgtggtgtc ctggttcgac cagggaacgc gggatgtgat cgggttgcgg 180

atcgccggtg gcgccatcct gtgggcgaca cccgatcaca aggtgctgac agagtacggc 240

tggcgtgccg ccggggaact ccgcaaggga gacagggtgg cgcaaccgcg acgcttcgat 300

ggattcggtg acagtgcgcc gattccggcg cgcgtgcagg cgctcgcgga tgccctggat 360

gacaaattcc tgcacgacat gctggcggaa gaactccgct attccgtgat ccgagaagtg 420

ctgccaacgc ggcgggcacg aacgttcggc ctcgaggtcg aggaactgca caccctcgtc 480

gccgaagggg ttgttgtaca caac 504

<210> 2

<211> 168
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
mini-intein derived from the full-length Mycobacterium
tuberculosis recA intein sequence

<400> 2
Ala Leu Ala Glu Gly Thr Arg Ile Phe Asp Pro Val Thr Gly Thr Thr
1 5 10 15

His Arg Ile Glu Asp Val Val Gly Gly Arg Lys Pro Ile His Val Val
20 25 30

Ala Ala Ala Lys Asp Gly Thr Leu His Ala Arg Pro Val Val Ser Trp
35 40 45

Phe Asp Gln Gly Thr Arg Asp Val Ile Gly Leu Arg Ile Ala Gly Gly
50 55 60

Ala Ile Leu Trp Ala Thr Pro Asp His Lys Val Leu Thr Glu Tyr Gly
65 70 75 80

Trp Arg Ala Ala Gly Glu Leu Arg Lys Gly Asp Arg Val Ala Gln Pro
85 90 95

Arg Arg Phe Asp Gly Phe Gly Asp Ser Ala Pro Ile Pro Ala Arg Val
100 105 110

Gln Ala Leu Ala Asp Ala Leu Asp Asp Lys Phe Leu His Asp Met Leu
115 120 125

Ala Glu Glu Leu Arg Tyr Ser Val Ile Arg Glu Val Leu Pro Thr Arg
130 135 140

Arg Ala Arg Thr Phe Gly Leu Glu Val Glu Glu Leu His Thr Leu Val
145 150 155 160

Ala Glu Gly Val Val Val His Asn
165

<210> 3
<211> 504
<212> DNA

<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
TOPO modified intein sequence from SEQ ID NO: 1

<400> 3
gccctcgcag agggcactcg gatcttcgat ccggtcaccg gtacaacgca tcgcatcgag 60

gatgttgtcg gtgggcgcaa gcctattcat gtcgtggctg ctgccaagga cggaacgctg 120

catgcgcggc ccgtggtgtc ctggttcgac cagggaacgc gggatgtgat cgggttgcgg 180

atcgccggtg gcgccatcct gtgggcgaca cccgatcaca aggtgctgac agagtacggc 240

tggcgtgccg ccggggaact ccgcaaggga gacagggtgg cgcaaccgcg acgcttcgat 300

ggattcggtg acagtgcgcc gattccggcg cgcgtgcagg cgctcgcgga tgccctggat 360

gacaaattcc tgcacgacat gctggcggaa gaactccgct attccgtgat ccgagaagtg 420

ctgccaacgc ggcgggcacg aacgttcggc ctcgaggtcg aggaactgca caccctcgtc 480

gccgaagggg tccttgtaca caac 504

<210> 4
<211> 168
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
TOPO modified intein sequence from SEQ ID NO: 2

<400> 4
Ala Leu Ala Glu Gly Thr Arg Ile Phe Asp Pro Val Thr Gly Thr Thr
1 5 10 15

His Arg Ile Glu Asp Val Val Gly Gly Arg Lys Pro Ile His Val Val
20 25 30

Ala Ala Ala Lys Asp Gly Thr Leu His Ala Arg Pro Val Val Ser Trp
35 40 45

Phe Asp Gln Gly Thr Arg Asp Val Ile Gly Leu Arg Ile Ala Gly Gly
50 55 60

Ala Ile Leu Trp Ala Thr Pro Asp His Lys Val Leu Thr Glu Tyr Gly
65 70 75 80

Trp Arg Ala Ala Gly Glu Leu Arg Lys Gly Asp Arg Val Ala Gln Pro

85 90 95

Arg Arg Phe Asp Gly Phe Gly Asp Ser Ala Pro Ile Pro Ala Arg Val
100 105 110

Gln Ala Leu Ala Asp Ala Leu Asp Asp Lys Phe Leu His Asp Met Leu
115 120 125

Ala Glu Glu Leu Arg Tyr Ser Val Ile Arg Glu Val Leu Pro Thr Arg
130 135 140

Arg Ala Arg Thr Phe Gly Leu Glu Val Glu Glu Leu His Thr Leu Val
145 150 155 160

Ala Glu Gly Val Leu Val His Asn
165

<210> 5
<211> 531
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
GATEWAY modified intein sequence from SEQ ID NO: 1

<400> 5
gccctcgcag agggcactcg gatcttcgat ccggtcaccg gtacaacgca tcgcatcgag 60

gatgttgtcg gtgggcgcaa gcctattcat gtcgtggctg ctgccaagga cggaacgctg 120

catgcgcggc ccgtggtgtc ctggttcgac cagggaacgc gggatgtgat cgggttgcgg 180

atcgccggtg gcgccatcct gtgggcgaca cccgatcaca aggtgctgac agagtacggc 240

tggcgtgccg ccggggaact ccgcaaggga gacagggtgg cgcaaccgcg acgcttcgat 300

ggattcggtg acagtgcgcc gattccgaca agtttgtaca aaaaagcagg cagcgcgcgc 360

gtgcaggcgc tcgcggatgc cctggatgac aaattcctgc acgacatgct ggcggaagaa 420

ctccgctatt ccgtgatccg agaagtgctg ccaacgcggc gggcacgaac gttcggcctc 480

gaggtcgagg aactgcacac cctcgtcgcc gaaggggttg ttgtacacaa c 531

<210> 6
<211> 177
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
GATEWAY modified intein sequence from SEQ ID NO: 2

<400> 6
Ala Leu Ala Glu Gly Thr Arg Ile Phe Asp Pro Val Thr Gly Thr Thr
1 5 10 15

His Arg Ile Glu Asp Val Val Gly Gly Arg Lys Pro Ile His Val Val
20 25 30

Ala Ala Ala Lys Asp Gly Thr Leu His Ala Arg Pro Val Val Ser Trp
35 40 45

Phe Asp Gln Gly Thr Arg Asp Val Ile Gly Leu Arg Ile Ala Gly Gly
50 55 60

Ala Ile Leu Trp Ala Thr Pro Asp His Lys Val Leu Thr Glu Tyr Gly
65 70 75 80

Trp Arg Ala Ala Gly Glu Leu Arg Lys Gly Asp Arg Val Ala Gln Pro
85 90 95

Arg Arg Phe Asp Gly Phe Gly Asp Ser Ala Pro Ile Pro Thr Ser Leu
100 105 110

Tyr Lys Lys Ala Gly Ser Ala Arg Val Gln Ala Leu Ala Asp Ala Leu
115 120 125

Asp Asp Lys Phe Leu His Asp Met Leu Ala Glu Glu Leu Arg Tyr Ser
130 135 140

Val Ile Arg Glu Val Leu Pro Thr Arg Arg Ala Arg Thr Phe Gly Leu
145 150 155 160

Glu Val Glu Glu Leu His Thr Leu Val Ala Glu Gly Val Val Val His
165 170 175

Asn

<210> 7
<211> 525
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
TOPO and GATEWAY modified intein sequence from
SEQ ID NO: 1

<400> 7
gccctcgcag agggcactcg gatcttcgat ccggtcaccg gtacaacgca tcgcatcgag 60

gatgttgtcg gtgggcgcaa gcctattcat gtcgtggctg ctgccaagga cggaacgctg 120

catgcgcggc ccgtggtgtc ctggttcgac cagggaacgc gggatgtgat cgggttgcgg 180

atcgccggtg gcgccatcct gtgggcgaca cccgatcaca aggtgctgac agagtacggc 240

tggcgtgccg ccggggaact ccgcaaggga gacagggtgg cgcaaccgcg acgcttcgat 300

ggattcggtg acagtgcgcc gattccgaca agtttgtaca aaaaagcagg cagcgcgcgc 360

gtgcaggcgc tcgcggatgc cctggatgac aaattcctgc acgacatgct ggcggaagaa 420

ctccgctatt ccgtgatccg agaagtgctg ccaacgcggc gggcacgaac gttcggcctc 480

gaggaactgc acaccctcgt cgccgaaggg gtccttgtac acaac 525

<210> 8
<211> 177
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
TOPO and GATEWAY modified intein sequence from
SEQ ID NO: 2

<400> 8
Ala Leu Ala Glu Gly Thr Arg Ile Phe Asp Pro Val Thr Gly Thr Thr
1 5 10 15

His Arg Ile Glu Asp Val Val Gly Gly Arg Lys Pro Ile His Val Val
20 25 30

Ala Ala Ala Lys Asp Gly Thr Leu His Ala Arg Pro Val Val Ser Trp
35 40 45

Phe Asp Gln Gly Thr Arg Asp Val Ile Gly Leu Arg Ile Ala Gly Gly
50 55 60

Ala Ile Leu Trp Ala Thr Pro Asp His Lys Val Leu Thr Glu Tyr Gly
65 70 75 80

```
Trp Arg Ala Ala Gly Glu Leu Arg Lys Gly Asp Arg Val Ala Gln Pro
                85                  90                  95

Arg Arg Phe Asp Gly Phe Gly Asp Ser Ala Pro Ile Pro Thr Ser Leu
            100                 105                 110

Tyr Lys Lys Ala Gly Ser Ala Arg Val Gln Ala Leu Ala Asp Ala Leu
        115                 120                 125

Asp Asp Lys Phe Leu His Asp Met Leu Ala Glu Glu Leu Arg Tyr Ser
    130                 135                 140

Val Ile Arg Glu Val Leu Pro Thr Arg Arg Ala Arg Thr Phe Gly Leu
145                 150                 155                 160

Glu Val Glu Glu Leu His Thr Leu Val Ala Glu Gly Val Leu Val His
                165                 170                 175

Asn
```

<210> 9
<211> 6262
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Synthetic
pET-GWMIT sequence

<220>
<221> modified_base
<222> (817)..(820)
<223> a, c, g, t, unknown or other

<220>
<221> modified_base
<222> (822)..(822)
<223> Inosine

<220>
<221> modified_base
<222> (823)..(826)
<223> a, c, g, t, unknown or other

<220>
<221> modified_base
<222> (827)..(828)
<223> Inosine

<220>
<221> modified_base
<222> (829)..(841)
<223> a, c, g, t, unknown or other

<400> 9
caaggagatg gcgcccaaca gtcccccggc cacggggcct gccaccatac ccacgccgaa 60

acaagcgctc atgagcccga agtggcgagc ccgatcttcc ccatcggtga tgtcggcgat 120

ataggcgcca gcaaccgcac ctgtggcgcc ggtgatgccg gccacgatgc gtccggcgta 180

gaggatcgag atctcgatcc cgcgaaatta atacgactca ctatagggga attgtgagcg 240

gataacaatt cccctctaga aataattttg tttaacttta agaaggaatt cgccctcgca 300

gagggcactc ggatcttcga tccggtcacc ggtacaacgc atcgcatcga ggatgttgtc 360

ggtgggcgca agcctattca tgtcgtggct gctgccaagg acggaacgct gcatgcgcgg 420

cccgtggtgt cctggttcga ccagggaacg cgggatgtga tcgggttgcg gatcgccggt 480

ggcgccatcc tgtgggcgac acccgatcac aaggtgctga cagagtacgg ctggcgtgcc 540

gccggggaac tccgcaaggg agacagggtg gcgcaaccgc gacgcttcga tggattcggt 600

gacagtgcgc cgattccgac aagtttgtac aaaaaagcag gcagcgcgcg cgtgcaggcg 660

ctcgcggatg ccctggatga caaattcctg cacgacatgc tggcggaaga actccgctat 720

tccgtgatcc gagaagtgct gccaacgcgg cgggcacgaa cgttcggcct cgaggaactg 780

cacaccctcg tcgccgaagg ggttcttgta cacaacnnnn mnnnnnnnnn nnnnnnnnnn 840

nggtaagcct atccctaacc ctctcctcgg tctcgattct acgcgtaccg gtcatcatca 900

ccatcaccat tgagtttgat ccggctgcta acaaagcccg aaaggaagct gagttggctg 960

ctgccaccgc tgagcaataa ctagcataac cccttggggc ctctaaacgg gtcttgaggg 1020

gttttttgct gaaaggagga actatatccg gatatcccgc aagaggcccg gcagtaccgg 1080

cataaccaag cctatgccta cagcatccag ggtgacggtg ccgaggatga cgatgagcgc 1140

attgttagat ttcatacacg gtgcctgact gcgttagcaa tttaactgtg ataaactacc 1200

gcattaaagc ttatcgatga taagctgtca aacatgagaa ttaattcttg aagacgaaag 1260

ggcctcgtga tacgcctatt tttataggtt aatgtcatga taataatggt ttcttagacg 1320

tcaggtggca cttttcgggg aaatgtgcgc ggaaccccta tttgtttatt tttctaaata 1380

cattcaaata tgtatccgct catgagacaa taaccctgat aaatgcttca ataatattga 1440

aaaaggaaga gtatgagtat tcaacatttc cgtgtcgccc ttattcccttt ttttgcggca 1500

ttttgccttc ctgtttttgc tcacccagaa acgctggtga aagtaaaaga tgctgaagat 1560

cagttgggtg cacgagtggg ttacatcgaa ctggatctca acagcggtaa gatccttgag 1620

agttttcgcc ccgaagaacg ttttccaatg atgagcactt ttaaagttct gctatgtggc 1680

gcggtattat cccgtgttga cgccgggcaa gagcaactcg gtcgccgcat acactattct 1740

cagaatgact tggttgagta ctcaccagtc acagaaaagc atcttacgga tggcatgaca 1800

gtaagagaat tatgcagtgc tgccataacc atgagtgata acactgcggc caacttactt 1860

ctgacaacga tcggaggacc gaaggagcta accgcttttt tgcacaacat gggggatcat 1920

gtaactcgcc ttgatcgttg ggaaccggag ctgaatgaag ccataccaaa cgacgagcgt 1980

gacaccacga tgcctgcagc aatggcaaca acgttgcgca aactattaac tggcgaacta 2040

cttactctag cttcccggca acaattaata gactggatgg aggcggataa agttgcagga 2100

ccacttctgc gctcggccct tccggctggc tggtttattg ctgataaatc tggagccggt 2160

gagcgtgggt ctcgcggtat cattgcagca ctggggccag atggtaagcc ctcccgtatc 2220

gtagttatct acacgacggg gagtcaggca actatggatg aacgaaatag acagatcgct 2280

gagataggtg cctcactgat taagcattgg taactgtcag accaagttta ctcatatata 2340

ctttagattg atttaaaact tcatttttaa tttaaaagga tctaggtgaa gatccttttt 2400

gataatctca tgaccaaaat cccttaacgt gagttttcgt tccactgagc gtcagacccc 2460

gtagaaaaga tcaaaggatc ttcttgagat cctttttttc tgcgcgtaat ctgctgcttg 2520

caaacaaaaa aaccaccgct accagcggtg gtttgtttgc cggatcaaga gctaccaact 2580

ctttttccga aggtaactgg cttcagcaga gcgcagatac caaatactgt ccttctagtg 2640

tagccgtagt taggccacca cttcaagaac tctgtagcac cgcctacata cctcgctctg 2700

ctaatcctgt taccagtggc tgctgccagt ggcgataagt cgtgtcttac cgggttggac 2760

tcaagacgat agttaccgga taaggcgcag cggtcgggct gaacgggggg ttcgtgcaca 2820

cagcccagct tggagcgaac gacctacacc gaactgagat acctacagcg tgagctatga 2880

gaaagcgcca cgcttcccga agggagaaag gcggacaggt atccggtaag cggcagggtc 2940

ggaacaggag agcgcacgag ggagcttcca gggggaaacg cctggtatct ttatagtcct 3000

gtcgggtttc gccacctctg acttgagcgt cgatttttgt gatgctcgtc agggggggcgg 3060

agcctatgga aaaacgccag caacgcggcc tttttacggt tcctggcctt ttgctggcct 3120

tttgctcaca tgttctttcc tgcgttatcc cctgattctg tggataaccg tattaccgcc 3180

tttgagtgag ctgataccgc tcgccgcagc cgaacgaccg agcgcagcga gtcagtgagc 3240

gaggaagcgg aagagcgcct gatgcggtat tttctcctta cgcatctgtg cggtatttca 3300

caccgcaatg gtgcactctc agtacaatct gctctgatgc cgcatagtta agccagtata 3360

cactccgcta tcgctacgtg actgggtcat ggctgcgccc cgacacccgc caacacccgc 3420

tgacgcgccc tgacgggctt gtctgctccc ggcatccgct tacagacaag ctgtgaccgt 3480

ctccgggagc tgcatgtgtc agaggttttc accgtcatca ccgaaacgcg cgaggcagct 3540

gcggtaaagc tcatcagcgt ggtcgtgaag cgattcacag atgtctgcct gttcatccgc 3600

gtccagctcg ttgagtttct ccagaagcgt taatgtctgg cttctgataa agcgggccat 3660

gttaagggcg gttttttcct gtttggtcac tgatgcctcc gtgtaagggg gatttctgtt 3720

catgggggta atgataccga tgaaacgaga gaggatgctc acgatacggg ttactgatga 3780

tgaacatgcc cggttactgg aacgttgtga gggtaaacaa ctggcggtat ggatgcggcg 3840

ggaccagaga aaaatcactc agggtcaatg ccagcgcttc gttaatacag atgtaggtgt 3900

tccacagggt agccagcagc atcctgcgat gcagatccgg aacataatgg tgcagggcgc 3960

tgacttccgc gtttccagac tttacgaaac acggaaaccg aagaccattc atgttgttgc 4020

tcaggtcgca gacgttttgc agcagcagtc gcttcacgtt cgctcgcgta tcggtgattc 4080

attctgctaa ccagtaaggc aaccccgcca gcctagccgg gtcctcaacg acaggagcac 4140

gatcatgcgc acccgtggcc aggacccaac gctgcccgag atgcgccgcg tgcggctgct 4200

ggagatggcg gacgcgatgg atatgttctg ccaagggttg gtttgcgcat tcacagttct 4260

ccgcaagaat tgattggctc caattcttgg agtggtgaat ccgttagcga ggtgccgccg 4320

gcttccattc aggtcgaggt ggcccggctc catgcaccgc gacgcaacgc ggggaggcag 4380

acaaggtata gggcggcgcc tacaatccat gccaacccgt tccatgtgct cgccgaggcg 4440

gcataaatcg ccgtgacgat cagcggtcca atgatcgaag ttaggctggt aagagccgcg 4500

agcgatcctt gaagctgtcc ctgatggtcg tcatctacct gcctggacag catggcctgc 4560

aacgcgggca tcccgatgcc gccggaagcg agaagaatca taatggggaa ggccatccag 4620

cctcgcgtcg cgaacgccag caagacgtag cccagcgcgt cggccgccat gccggcgata 4680

atggcctgct tctcgccgaa acgtttggtg gcgggaccag tgacgaaggc ttgagcgagg 4740

gcgtgcaaga ttccgaatac cgcaagcgac aggccgatca tcgtcgcgct ccagcgaaag 4800

cggtcctcgc cgaaaatgac ccagagcgct gccggcacct gtcctacgag ttgcatgata 4860

aagaagacag tcataagtgc ggcgacgata gtcatgcccc gcgcccaccg gaaggagctg 4920

```
actgggttga aggctctcaa gggcatcggt cgagatcccg gtgcctaatg agtgagctaa   4980

cttacattaa ttgcgttgcg ctcactgccc gctttccagt cgggaaacct gtcgtgccag   5040

ctgcattaat gaatcggcca acgcgcgggg agaggcggtt tgcgtattgg gcgccagggt   5100

ggtttttctt ttcaccagtg agacgggcaa cagctgattg cccttcaccg cctggccctg   5160

agagagttgc agcaagcggt ccacgctggt ttgccccagc ag
```